

FIG. 1

Minimum signal peptide score	false positive rate	false negative rate	proba(0.1)	proba(0.2)
3.5	0.121	0.036	0.467	0.664
4	0.096	0.06	0.519	0.708
4.5	0.078	0.079	0.565	0.745
5	0.062	0.098	0.615	0.782
5.5	0.05	0.127	0.659	0.813
6	0.04	0.163	0.694	0.836
6.5	0.033	0.202	0.725	0.855
7	0.025	0.248	0.763	0.878
7.5	0.021	0.304	0.78	0.889
8	0.015	0.368	0.816	0.909
8.5	0.012	0.418	0.836	0.92
9	0.009	0.512	0.856	0.93
9.5	0.007	0.581	0.863	0.934
10	0.006	0.679	0.835	0.919

FIG. 2

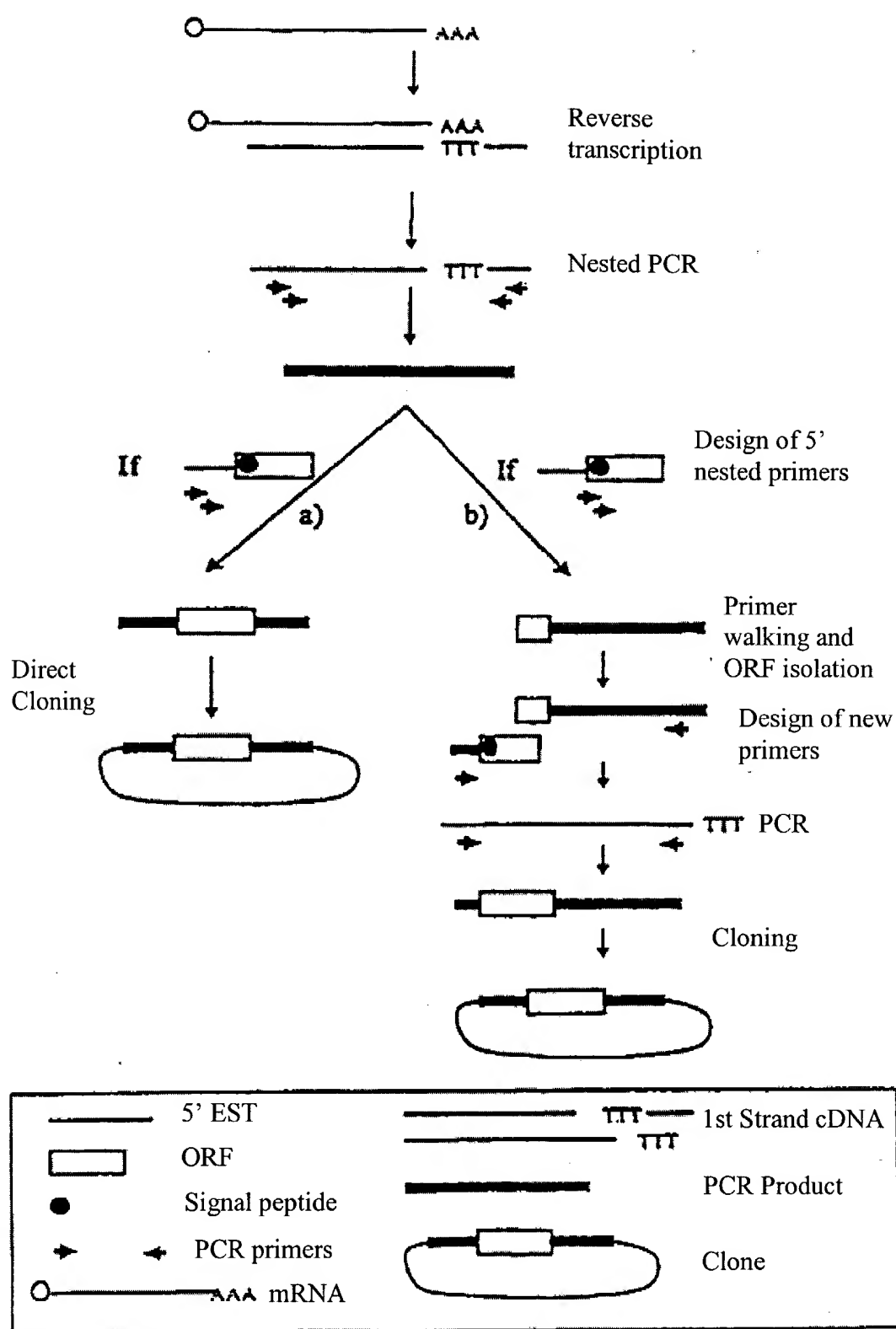


FIG. 3

Description of promoters structure isolated from SignalTag 5' ESTs

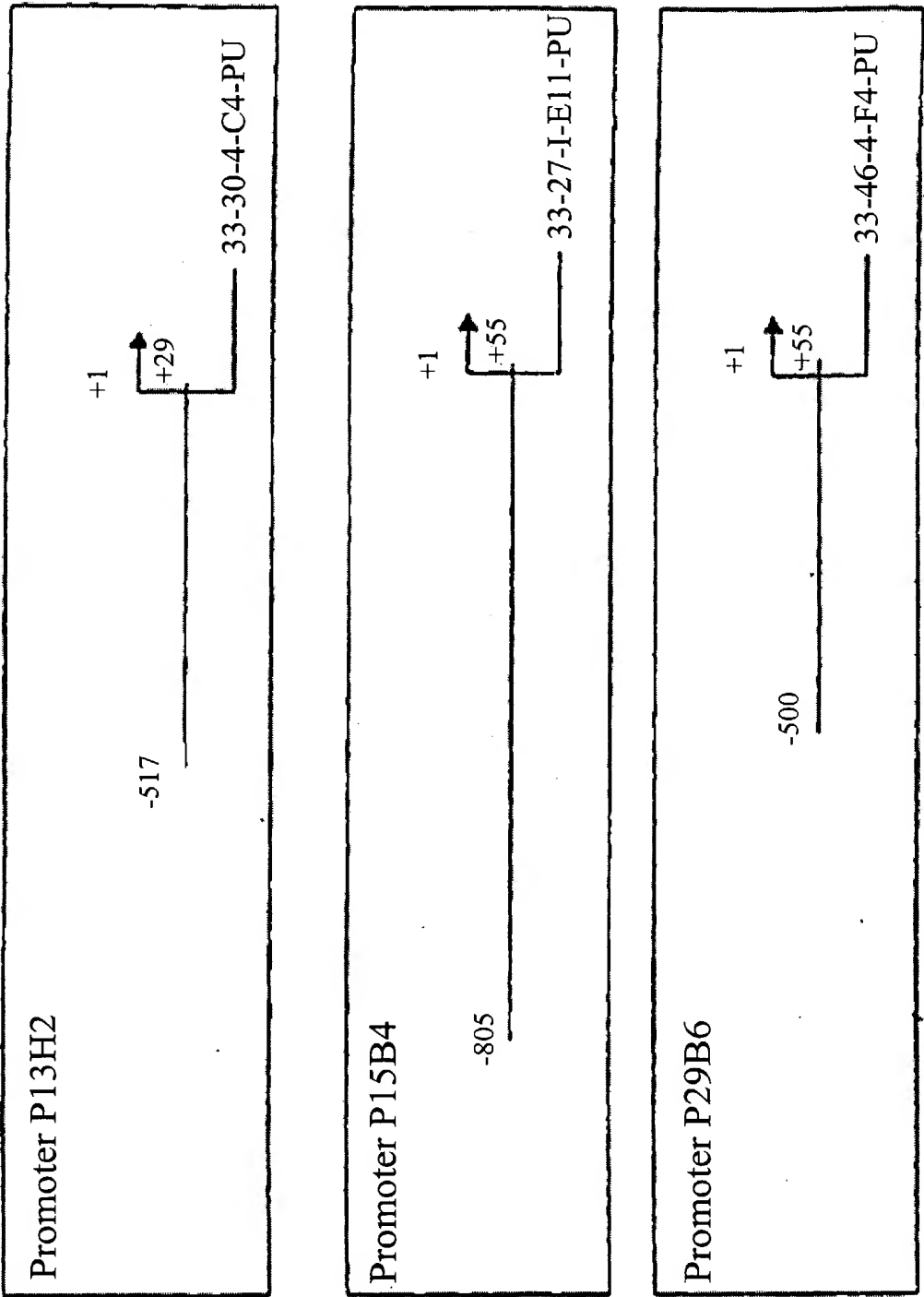


FIG. 4

**Description of Transcription Factor Binding Sites present on promoters isolated from SignalTag sequences**

**Promoter sequence P13H2 (546 bp):**

Matrix	Position	Orientation	Score	Length	Sequence	Location in:
						<b>SEQ ID NO: 17</b>
CMYB_01	-502	+	0.983	9	TGTCAGTTG	17-25
MYOD_Q6	-501	-	0.961	10	CCCAACTGAC	complement of 18-27
S8_01	-444	-	0.960	11	AATAGAAATTAG	complement of 75-85
S8_01	-425	+	0.966	11	AACTAAATTAG	94-104
DELTAEF1_01	-390	-	0.960	11	GCACACCTCAG	complement of 129-139
GATA_C	-364	-	0.964	11	AGATAAATCCA	complement of 155-165
CMYB_01	-349	+	0.958	9	CTTCAGTTG	170-178
GATAI_02	-343	+	0.959	14	TTGTAGATAGGACA	176-189
GATA_C	-339	+	0.953	11	AGATAGGACAT	180-190
TALIALPHAE47_01	-235	+	0.973	16	CATAACAGATGGTAAG	284-299
TALIBETA47_01	-235	+	0.983	16	CATAACAGATGGTAAG	284-299
TALIBETAITF2_01	-235	+	0.978	16	CATAACAGATGGTAAG	284-299
MYOD_Q6	-232	-	0.954	10	ACCATCTGTT	complement of 287-296
GATAI_04	-217	-	0.953	13	TCAAGATAAAGTA	complement of 302-314
IK1_01	-126	+	0.963	13	AGTTGGGAATTCC	393-405
IK2_01	-126	+	0.985	12	AGTTGGGAATTCC	393-404
CREL_01	-123	+	0.962	10	TGGGAATTCC	396-405
GATAI_02	-96	+	0.950	14	TCAGTGATATGGCA	423-436
SRY_02	-41	-	0.951	12	TAAACAAAACA	complement of 478-489
E2F_02	-33	+	0.957	8	TTTAGCGC	486-493
MZF1_01	-5	-	0.975	8	TGAGGGGA	complement of 514-521

**Promoter sequence P15B4 (861 bp):**

Matrix	Position	Orientation	Score	Length	Sequence	Location in:
						<b>SEQ ID NO: 20</b>
NFY_Q6	-748	-	0.956	11	GGACCAATCAT	complement of 60-70
MZF1_01	-738	+	0.962	8	CCTGGGGA	70-77
CMYB_01	-684	+	0.994	9	TGACCGTTG	124-132
VMYB_02	-682	-	0.985	9	TCCAACGGT	complement of 126-134
STAT_01	-673	+	0.968	9	TTCTTGGA	135-143
STAT_01	-673	-	0.951	9	TTCCAGGAA	complement of 135-143
MZF1_01	-556	-	0.956	8	TTGGGGGA	complement of 252-259
IK2_01	-451	+	0.965	12	GAATGGGATTTT	357-368
MZF1_01	-424	+	0.986	8	AGAGGGGA	384-391
SRY_02	-398	-	0.955	12	GAACAAAACA	complement of 410-421
MZF1_01	-216	+	0.960	8	GAAGGGGA	592-599
MYOD_Q6	-190	+	0.981	10	AGCATCTGCC	618-627
DELTAEF1_01	-176	+	0.958	11	TCCACCTTCC	632-642
S8_01	5	-	0.992	11	GAGGCAATTAT	complement of 813-823
MZF1_01	16	-	0.986	8	AGAGGGGA	complement of 824-831

**Promoter sequence P29B6 (555 bp):**

Matrix	Position	Orientation	Score	Length	Sequence	Location in:
						<b>SEQ ID NO: 23</b>
ARNT_01	-311	+	0.964	16	GGACTCACGTGCTGCT	191-206
NMYC_01	-309	+	0.965	12	ACTCACGTGCTG	193-204
USF_01	-309	+	0.985	12	ACTCACGTGCTG	193-204
USF_01	-309	-	0.985	12	CAGCACGTGAGT	complement of 193-204
NMYC_01	-309	-	0.956	12	CAGCACGTGAGT	complement of 193-204
MYCMAX_02	-309	-	0.972	12	CAGCACGTGAGT	complement of 193-204
USF_C	-307	+	0.997	8	TCACGTGC	195-202
USF_C	-307	-	0.991	8	GCACGTGA	complement of 195-202
MZF1_01	-292	-	0.968	8	CATGGGGA	complement of 210-217
ELK1_02	-105	+	0.963	14	CTCTCCGGAAGCCT	397-410
CETSIP54_01	-102	+	0.974	10	TCCGGAAGCC	400-409
API_Q4	-42	-	0.963	11	AGTGACTGAAC	complement of 460-470
APIF_Q2	-42	-	0.961	11	AGTGACTGAAC	complement of 460-470
PADS_C	45	+	1.000	9	TGTGGTCTC	547-555

**FIG. 5**